

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2001, 18:33:32 ; Search time 6385.92 Seconds  
(without alignments)  
8930.521 Million cell updates/sec

Title: US-09-117-447-1  
Perfect score: 3687  
Sequence: 1 atgataagaaagaaagctgt.....ttacatctcatataagtaa 3687

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_pn:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_hcgo\_hum:\*  
20: em\_hcgo\_inv:\*  
21: em\_hcgo\_rod:\*  
22: em\_hcg\_hum1:\*  
23: em\_hcg\_hum2:\*  
24: em\_hcg\_hum3:\*  
25: em\_hcg\_hum4:\*  
26: em\_hcg\_hum5:\*  
27: em\_hcg\_hum6:\*  
28: em\_hcg\_hum7:\*  
29: em\_hcg\_hum8:\*  
30: em\_hcg\_inv1:\*  
31: em\_hcg\_inv2:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_pn:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_hcg1:\*  
61: gb\_hcg2:\*  
62: gb\_hcg3:\*  
63: gb\_hcg4:\*  
64: gb\_hcg5:\*  
65: gb\_hcg6:\*  
66: gb\_hcg7:\*  
67: gb\_hcg8:\*  
68: gb\_hcg9:\*  
69: gb\_hcg10:\*  
70: gb\_hcg11:\*  
71: gb\_hcg12:\*  
72: gb\_hcg13:\*  
73: gb\_hcg14:\*  
74: gb\_hcg15:\*  
75: gb\_hcg16:\*  
76: gb\_hcg17:\*  
77: gb\_hcg18:\*  
78: gb\_hcg19:\*  
79: gb\_hcg20:\*  
80: gb\_hcg21:\*  
81: gb\_hcg22:\*  
82: gb\_hcg23:\*  
83: gb\_hcg24:\*  
84: gb\_hcg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_rot1:\*  
95: gb\_rot2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	100.0	3687	9	A93832
2	3687	100.0	3687	9	AX000218
3	3687	100.0	3687	9	AX033303
4	3687	100.0	3687	2	BSSBSA
5	513	13.9	3680	1	AF055578
6	487.6	13.2	3090	2	AF228338
7	77.4	2.1	2069	45	E10125
8	77.4	2.1	3399	45	E10126

A93832 Sequence 1  
AX000218 Sequence  
AX033303 Sequence  
X71092 B-stearothe  
AF055578 Bacillus  
AF228338 Bacillus  
E10125 DNA encodin  
E10126 DNA encodin

[illegible]

## ALIGNMENTS

[illegible]

```

LOCUS           A93832               3687 bp    DNA
DEFINITION      Sequence 1 from Patent WO9728263.
ACCESSION       A93832
VERSION         A93832.1
KEYWORDS        GI:6741963
SOURCE           Bacillus stearothermophilus.
ORGANISM        Bacillus stearothermophilus
                 Bacteria; Firmicutes; Bacillus/Clostridium group;
                 Bacillus/Staphylococcus group; Bacillus.
REFERENCE       1 (bases 1 to 3687)
AUTHORS         Lubitz,W. and Sleytr,U.
TITLE           RECOMBINANT EXPRESSION OF S-LAYER PROTEINS
JOURNAL         Patent: WO 9728263-A 1 07-AUG-1997;
                 LUBITZ WERNER (AT); SLEYTR UWE (AT)
FEATURES        Location/Qualifiers
                 1..3687
                 /organism="Bacillus stearothermophilus"
                 /strain="PV72"
                 /db_xref="taxon:1422"
                 /clone="SBSA"
                 1..50
                 1..3687
                 /note="unnamed protein product"
                 /codon_start=1
                 /transl_table=1
                 /protein_id="CAB69613.1"
                 /db_xref="GI:6741964"
                 /translation="MDKKKAVIKLATASAIASAFAVANPNNSAEATVAVTVSQAKQ
                 EKRAAYVSHTVETGEFFPNINOVYAEYNKRRKYDPAVALYNNAGGAKKDAYLADIQ

```

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

[illegible]

Dh	1741	GAAATATCAATCTATCAAAATTAATAAAAAAGCTTGTAAGTCCCATTAAGATTGAA	1800
Qy	1801	ttaggcaatgcttaacggaacaacatlgatltcaaaactcaagacttaactgctcciaa	1860
Dh	1801	TTTAGGCACTGGTTAACGAGAAACATPATGATGTTCAAAATCTAGACTTCTCTTCA	1860
Qy	1861	gtattatcgtaacgctctaaaatgycgcgcgtlgatltaaeaqtaactgaaqccaaga	1920
Dh	1861	GTTATTACCGCTTACGCTCTMAAAATGCGAGCGCTGGATTAAAGTAACTGAACTCAAGAA	1920
Qy	1921	tttactggaagctcttaaggaatttaaatataatltgaatgctacaacogtltccggtagc	1980
Dh	1921	TTTACTGTGAAAGCTTCAGAGAAATTTAAATACATTTAAATGCTACAAACCGCTTTCGGGTAGC	1980
Qy	1981	acaatcaacatcggctcaagttgctgtagtaaaqccgggtgycaaactaactgctctta	2040
Dh	1981	ACAAATCAATACGGTTCAGATTTGCTGTGATTAAGCGCGGTGCAAACTTATCTGCTCTTCA	2040
Qy	2041	gcaagtgacacaceltccagcttagtltgaanqcgtaactgctcaaga tggaaacatacaa	2100
Dh	2041	GCAAGTGACATCATTTCCAGCTAGGTGAAAGCGGTTACTGTCTACAGATGGAACTATCAAA	2100
Qy	2101	gtgaagttgctgcttaaccaaattagaacgylaaccaaggttacaactagtagtgcgt	2160
Dh	2101	GTTGAAAGTTGCTGCTTAACCAATTATGAAAGTAAACGAAGGTTACAAATTAATAGTCTTCCGT	2160
Qy	2161	aaagtgcaaacgctctgtttaaagtgtctgcaaatgycaaatlaactlttagcaactaact	2220
Dh	2161	AAAGTGCACACGCTCTGTTAAAGATGCTGCAAATAGCAAAATATCTTTACCACTATACAT	2220
Qy	2221	actataactlttaacacactgaaaggltcaagacgtlaacagccacacaggtltacaaagtlctc	2280
Dh	2221	ACTATACATTTTAACTGAAAGTCAAGAGCTAACAGCAACCAACGGTTACAAAGCTATTCC	2280
Qy	2281	aaagtgatctcttlaaaagccgctgtgcagttaactacacttaagaaogtltga tgcagct	2340
Dh	2281	AAAGTGATCTCTTAAAGAGACGCTGATGCTACTTACACTTACGAACTGATATCAGCT	2340
Qy	2341	caaaatltcaactacgaatttagcgaagaatttaaaacttcaagtgtttcttttagtggct	2400
Dh	2341	CAAAATTCACATTCACAAATTTAGCGAAGAAATTAATAACTTCTAGTGGTCTTTACTGGCT	2400
Qy	2401	ggcaagatgaactgttcgaagaatttaacaaacaacgcgtgtggtagatgctgtactggaaca	2460
Dh	2401	GGCAAGATGAACGTTCGAGAAATTAACAAACAAACGGAATGGGTGATGCTGTACTGGAAACA	2460
Qy	2461	actgtatcagttgctctcctaaagacagatgycaaatgyltaaaqtaacacagctltgctgtaca	2520
Dh	2461	ACTGTATCAGTTGCTCTTAAGACAGATGCAAAATGTAAGTTAAACAGCTTCTGTGGTTACA	2520
Qy	2521	ttaaactgtctctgtgccaataacgcgaagaagatgycgaaaatltgcgtcgtgtagtagataagctc	2580
Dh	2521	TTAACTGCTCTTGCACATTAACGACAAAGATGCGAAATTTGCTGCTGCTATTAATTAAGCT	2580
Qy	2581	tciaactgtatgaaatltgtgtgtgtgactgctggtgaatgtaacttaagaaaaagatlttaact	2640
Dh	2581	TCTACTGATGGAATTCTGTAATGAACTGGTAATGTAATTAAGGAAAAACATATTTAATT	2640
Qy	2641	cgtlatacaacagctgtagacacacactgttagctctgtgaaagcgtgcgtgylgacaagaatgylt	2700
Dh	2641	CGTTATCAAAACGCTGACACACACTGTATAGCTTGTGTGAAAGCTGCTGCTACAAACATGCT	2700
Qy	2701	caaaagcgtctgtgtgtatctcccaaaagacgtgcgaattgtatacaaaccaagagcttata	2760
Dh	2701	CAAAAGCCTTCTGTGATTTCCCAACACACACTGCAATTTGAATACAAATTAAGAGCTTAATTA	2760
Qy	2761	gttgaatcaatgaaactgtatttagcggagatlaaacctggaacacatcgttgtttaaagat	2820
Dh	2761	GTTGAATTCAAATGAAACTGAATTTAGCGGAAGTTAAACCTGAGAAACATGCTTGTTAAGAAAT	2820
Qy	2821	gcacagcgttaatgcgtgtagctgtgtactgtlaacagcatltagacggtltctacaataaaatt	2880

[illegible]

EXTRACELLULAR SPACE  
Patent: WO 9906567-A 1 11-FEB-1999;  
LUBRIT WERNER (AT); RESCH STEPHANIE (DE)  
FEATURES  
source  
1. .3687  
/organism="Bacillus stearothermophilus"  
/strain="pv72"  
/db\_xref="taxon:1422"  
/clone="SB5A"  
1. .90  
1. .3687  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAB77069.1"  
/db\_xref="gi:7240683"  
/translation="MDRRKAKVLTATASAIASAFAVAMPNASEADVATVVSQAKQ  
FKKAYVYSHVETVEGEPININDV A EYNNARK YRDAVALVYNAGGAKDAVLAIDQ  
KEVTVYKAMPKSGEARVAT YIDAVYVATLDMROLEPAVQAKLEAGEYHKI  
PYELKTRFVILLDRYGGKTRDILKSTPKAKAOELRDSLIYDIPYAKARVQDAVQK  
MLDKAAAOINOYILPKVTDAPKTEL TEVAKAOLDEALPIHPVYSVAITONKKA  
VELTAVPVGKTLKQLLSAANEDTVNNYVITIKVDENITPALNADVLSSTGKRTT  
VDASTPEENETVEYVVGKGLDKNGKEFKEDALFEKLRNDVAVTQVGVNTNFSYN  
LAQTPQTDQTLTVFEPKILAPETVNSNVTIDVEGKRPIYASTGSGTITITLKE  
ALVYGRQYKLA INNKTLTGNAVAVELVTANASAPTVAPTLTGSTLTGSLTT  
NMVKLAGVNEAGTYPPGLQPTTFATKIDESTLADNPFLVERSGTVASSELKYNA  
DAKWTLPVRADLKENTIYQIKIKKLSQKGIELGVNKTVEFKQDILAPVLSV  
TSKQDAGIKVTPAEQETVYKSENLATPNATVYSGSNITYGVAVYVQVMSALTYS  
DIDPASYEATVGDQGTKYKVAANOJERNRGYKLVYGRKATPVRDAAANTLATY  
IYVTEGQDVTAPTVKVRGDSKQADAVTLTNDVQOKFTIQSEELKTSGL  
VGVTVTEKLTNNGVMDAGTGYTVA PKTDANGKVTAAVYTLTGDNNDKKLRLV  
VDKSTGQIADVAGNVIKERDILRYMSRHVTVA SVAAADKDGNSAAPFSTAIID  
TTLKSLVEPNETDLAEVKEPENVYKDAAGNAVATVYAGDSNKEFVTSQELKAGT  
VYSTIDGVNRDVGNTISKYTSFKYTSAPLTSIS IAGAVNDSKRTITLESQS  
VSPNTILFKRADGSPNNYTLVYNNENKTKYKIFEHGATLDEFTQELVNSDFQF  
TDIDSKYTFIDGYSATDBVYKDALVGSNNGTSTYDQAAATRLRSADVAPRYAOLF  
SEGIDLTNAITVNTDDTKVEYIESVSDHDDHDAATKETLVINTVTLVLDNRKT  
YKIVVSGKDAAGVNAVDITVEYIK"  
91. .3684  
mat\_peptide 1316 a 660 c 715 g 996 t  
BASE COUNT 1316 a 660 c 715 g 996 t  
ORIGIN

Query Match 100.0%; Score 3687; DB 9; Length 3687;  
Best Local Similarity 100.0%; Pident. No. 0;  
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 atggtatgaaanaagcgtgtgaacttaacacagaactgtatattgacgcagtgcaatt 60  
1 ATGGATAGGAAANAAGCTGTGAACCTAGCACACAGCACTGCTATTGCGAAGCTGCATT 60  
61 gtccgtgcgaatcccaacgcctctcgaacgcgtctacagatgttagcacacgtgtgaagccaa 120  
61 GTCCGCTGCAAAATCCAAAGCGCTTCTAAGCGGCTACAGATGTAGCAACAGTATGTAAGCCAA 120  
121 gcaaaagcacaggttcaaaaaagcatataactataactacagcaatacagtaaacgynaacttgc 180  
121 GCAAAAGCACAGGTTCAAAAAAGCATATTACTTACACCCATACAGTAAACGGAACCTGGT 180  
141 gaattcccaaaaactaaagatgatatgtctatataacaaanaagggaaaaaagcgtacgt 240  
141 GAATTCCCAAAACATTAAAGATGTATATCTGTAATTCACAAAGCGAATAAAACGATACCGT 240  
241 gatcgcgcagcatctagttaatacaagcaggtgcgcgaaaaaagacgccttacttaagtgat 300  
241 GATCGCGTAGCATTAAGTATTAATTAACCAAGTGTGGCGGAAAAAAGACGCTTACTTAGCTGAT 300  
301 ttacaaaaagaatatgaacttaacgttttcaaaagcaaacccnaaactctgcgaaactcgt 360  
301 TTACAAAAAGAAATATGAACCTTACGTTTTCAAACCAACCTTAATCTGCGGAAGCTCGT 360  
361 gttagaacttaccatgcagcttcaaacatctgcacaacaaatlaagcgaatctgcgcaagag 420

Dh	361	GTGACAACTTACATCGATGCTTACAACTATGCACAAAAATTACAGCAAAATGCGCCAAAG	420
Qy	421	ctgaagctgctgtctcaagcaaaagattlagaagaagcacaacatctcaacaatc	480
Dh	421	CTGAGGCTGCTGTCAAGCAAAAGATTTCGAAAAAGCAGCAACTCTATCAGAAATT	480
Qy	481	ccttatgaattaaacclcgacagctcatcttagatcgctataagylaaacacactcgt	540
Dh	481	CCATTGAAATTAATAACTGCGACAGTCATTTTACATGCGGTATATGTTAAACAMCTGT	540
Qy	541	gattactctgcgtctctatctttaaagcaaaagcacaagaacttcgcgcagcttaattac	600
Dh	541	GATTATCTGCTCTCATTTTAAAGCAAAAGCAAGAACTTCGCGACACTTAATTAT	600
Qy	601	gatatlacgcttgcaatgaaagcgcgcaagtgaaagcaagcgtctgaaagagccaattc	660
Dh	601	GATATTACGCTTGCAATGAAAGCGCGGAAAGTACACAGCCCTGTGAAGCAAGCAATTTA	660
Qy	661	gacaaagctaaagctgctgttga tcaaatcaatcaatcttaccnaaagtaaacagctgc	720
Dh	661	GACAAAGCTAAAGCTGCTGTGATVCAAAATCAATCAATCTACCAAAAGTAAACAGATGCT	720
Qy	721	tccaactcgaactaacagaagtagcgaaaaaagcattagaatgcagatgcgtc	780
Dh	721	TTTCAAACTGAACTTAACAGAAAGTACGAAAAAGCATTAATGCAATGCAAGTCCGCTT	780
Qy	781	actccaagaagtgaaagtgaaatgcaatcaacccaacaaagaagtggttgaattaca	840
Dh	781	ACTCCAAAGTTGAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGT	840
Qy	841	gcagtlaccagtgaaagcaaacacaaatcaacacttcagctgcgcgaatgaaataga	900
Dh	841	GCAGTACCAAGTGAAGCAACACTAAATTAACAATCTTACCTGCAAAATGAAGTAAACA	900
Qy	901	gtaaaagtaataatcgtatgaatctataaagtggaagcaatccattcgccttaac	960
Dh	901	GTAAGGTAATTAATCTAGCTATCTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	960
Qy	961	acggcagatggttctctatctatacagaagcaaaactatcagcttgatggttcaactcca	1020
Dh	961	ACGGCAGATGTTCTTATCTATCTACAGAGCAAGCAAACTATCTGCAATGCTTCAACTCCA	1020
Qy	1021	ttccgaataatacaggaagtaaaagtagtagtlaaaggtltaaagaacaaatgagcaaa	1080
Dh	1021	TTCCGAAATTAATACGAGATTAATAGTATAGTATTAAGGTTTAAAGCAAAATTCGAAA	1080
Qy	1081	gaatttaaagaagatgcattcactcttcaagctcgaagaatgctgtagtactcaagt	1140
Dh	1081	GAAATTAAGAAAGAAATCACTTCTCAAGCTTCAAGTATGCTGTAACTCAAGTGT	1140
Qy	1141	tttggagacaaatgtaacaaacaaacactctgttaacttgacgaagtlacttcgcgact	1200
Dh	1141	TTTTGGAACTTAATGTAAACAAACAACTTCTGTAAACTTACAGACAGCTACTTTCGAACT	1200
Qy	1201	gaacgatacttlaacagtagtattgtaagtttgtagcaactgaaactgaaacagctgcg	1260
Dh	1201	GACGATACCTTAAACACTAGTATTTGATTAAGTTGTAGCAACCTCAAACTGTAACAGCTGC	1260
Qy	1261	aacgcttactataacagatgttgaactcgaacaaacgaatccagtaattgcatactactct	1320
Dh	1261	AACGTTACTATTACAACTATTTGAACCTGCAAAACGCAATTCACACTAATTTGCAATCTCT	1320
Qy	1321	ggcttcacaaatctactatcgctttaaagaagcggttagtaacggttgaacaaatgaactc	1380
Dh	1321	GCTTCTACAAATTAATTAATGCTTAAAAAGAGCGCTTAACTGTAACCTGTAACAAATTAACCT	1380
Qy	1381	gclatcaatatagtltaaacaatlaaactggtltaacaatgcgaagaagcttaacagctagctc	1440
Dh	1381	GCTATCAATTAATTAATGCTTAAAAAGAGCGCTTAACTGTAACCTGTAACAAATTAACCT	1440
Qy	1441	actgcaaacgcatcagaccacactgttgctacgcgtccactactcttaagtgtaacaact	1500
Dh	1441	ACTGCAAAAGCAATGCAACAACTGTTGCTACCGCTCTACTACTTAACTGCTGTAACCT	1500
Qy	1501	ttatctactgctctcttacaacaagaatgttgggtaaatctgagctggtgtaagta	1560
Dh	1501	TTATCTACTGCTCTCTTACAAACAAAGTTTGGGTAATATGCTGCTGCTGTAAGTAA	1560
Qy	1561	gcvgaaacttatactcgtgctcaatctcaacaaacagcttgctactaagltagaagaa	1620
Dh	1561	GCTGCAACTTATTAATCTGCTGCTCAATTAACAAACAGCTTGCCTAACTTAAGCAAG	1620
Qy	1621	ttcatcttagctgtaactttgtattglttgaaagaagaatctgttaagtggtgctctc	1680
Dh	1621	TTTACTTTAGCTGTAATCTTGTATGTTGAAAAAGAAATCTGTACAGTTGCTGCTCT	1680
Qy	1681	gaactaaataataatgcagacagcctaaatgttcaacttagtgcccaaaagcgagacttaa	1740
Dh	1681	GAACTAAATTAATTAATGCAAGCTTAATTAATGTTAACTTATGCTCCAAAAGCGACTTAA	1740
Qy	1741	gaaatacaatctatacaatcaaaatlaaaaaagccttgaaagtcgataaagtaagtaa	1800
Dh	1741	GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1800
Qy	1801	ttagagcctgtltaacgagaaacataagatltcaaaaactcaagacttaactgctccata	1860
Dh	1801	TTAGAGCTGCTTAAAGCAAAAGCAATGAGTCAAACTCAAGCTTAATGCTGCTTAA	1860
Qy	1861	gtttatagcgtlaaagctttaaagaagcgagcgttgatataaagtaaacgaaagcaga	1920
Dh	1861	GTTATTAAGCTTAAGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1920
Qy	1921	ttctacttgaaagctctcagagaaatcttaatacatcttaatgcttaaacgcgtctcgag	1980
Dh	1921	TTTACTGCTGAAGTCTCAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1980
Qy	1981	acaatcacaacagctcagatgtgctgtaaaagcgagtgcaaatctatctgctctaca	2040
Dh	1981	ACAATCAACATACGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2040
Qy	2041	gcaagtagacatctatcagagtagtggttgaaagcggttactcgttgaaagtgaaatacaa	2100
Dh	2041	GCAAGTAGCATCTATCTCAAGTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2100
Qy	2101	gtgaaagtgtcgtgtaacaaatctgaacagcgttaacaaaggttaacaaatctgagctgcgt	2160
Dh	2101	GTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2160
Qy	2161	aaaggtgcaacagctcctgtttaaagatgctgcaaaatgcgaatctttagcaaatcaact	2220
Dh	2161	AAAGGTGCAACAGCTCCTGTTAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2220
Qy	2221	atctatacatcttaacatctgaaggtcgaaggttaacagacaaacaggtttaaagaatctc	2280
Dh	2221	ATCTATCATTTTAACCTGAAGCTGCAAGCTTAACAGCTTAACAGCTTAACAGCTTAACAG	2280
Qy	2281	aaaggtatcttlltaaagaacgctgagctgtaactcaacttaacgaactgagctgagcag	2340
Dh	2281	AAAGGTATCTTAAAGAGCGCTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	23

```

OY 2581 tctactgataagattgctgtagtgaagcggtaagttaataagaanaagataatttaatt 2640
DB 2581 TCTACTGATGGAATTTGCTGATGAGCTGTAATGTAATTAAAGAAAAGTAATTTAAAT 2640
OY 2641 cgttataaagcttgagagaacacactgtagcttctgtgaagcgtgctgaacaaagatggt 2700
DB 2641 CGTTAAAGAGCTGGAGACACACTGTAGCTTCTGTGTAAGAGCTGCTGCTAACAAGATGCT 2700
OY 2701 caaaagccttcgctgcatctccaaacagcactgcaattgatacaataagagcttata 2760
DB 2701 CAAAGCGCTTCTGCTCATTTCCCAACAGACTGCATTTGATTCACAACTAAGACTTATTA 2760
OY 2761 gttagaattcaatgaacatgtaattagcgggaagttaaacctgagaaacatcgctgttaagaat 2820
DB 2761 GTTAGAATTCATGAACCATGATTTAGGGGAAGTAAACCTGAGAACATCGTTGTTAAAGAT 2820
OY 2821 ggaagcgggaatagcggtagctgtagtgaacagcactgaagcgttctcaataaatt 2880
DB 2821 GGAAGCGGAATAGCGGTAGCTGCTGTAACAGCATTTAGACGCTTCTCAATAATTAATTT 2880
OY 2881 gttatcacctccatctcaagaatlaaaagctggtagacgttactctgtaacaattgacgt 2940
DB 2881 GTTATCACCTCCATCTCAAGAAATTTAAAGCTGGTAGAGTTACTCTGTAAACAATTGACGT 2940
OY 2941 gtgagagataagtaggtagaacaacatctctaaatacattactcgttcaagactgtacct 3000
DB 2941 GTGAGAGATAAGTAGGTAAACACAAATCTCTAAATACATTACTGCTCAAGACTGTACTCT 3000
OY 3001 ggaagcagcagttactctcaatccagcatctgctgaagcgtgacgttgaacgttct 3060
DB 3001 GGAAGCAGCAGTTACTCTCAATCCAGCATTTGCTGACGGTGGCAAGTTAACGTTGACCGTCT 3060
OY 3061 aaaacatlaacatgaattcagcga tcaattccaaacccaacaa tcaactttaagaag 3120
DB 3061 AAAACATTAACATGAATTTCAAGCATTTGAGTTCCAAACCAACATCACTACTTTAAGAG 3120
OY 3121 gctgacggaactcattctcaatacactttaagtaaatgaataaayaaataaaca 3180
DB 3121 GCTGACGGAACTCAATTTACTAATTAACCTTTAGTAAATGTAATGTAATGTAATGTAATGTA 3180
OY 3181 tacaataatgtagtccaaagaagtgtaaacacttgaagtttaactaataatgaagtagca 3240
DB 3181 TACAATAATGTAATCCCAAGAGGTGAACACTTGACGAGCTTCACTCAATTAATGAGTTAAGA 3240
OY 3241 gttccaaagaatttccaaacgtgtagtgaattatgataagaagttcaattcacaggt 3300
DB 3241 GTTCCAAAGAATTTTCCAAACGTGACTGATTAATGATAGCAAAAGTTACATTCACACAGCT 3300
OY 3301 tctgttgcactgaagcgaagtaaaacctgctctagtagagcgttggctcatggaatggaaca 3360
DB 3301 TCTGTTGCTACTGACGAAGTAACCTGCTCTAGAGCGCTTGCTCATGAGTAATGGAACA 3360
OY 3361 agctatlaacagatgctgtagcagcagacagactcgtctgtagcttctgtggag 3420
DB 3361 AGCTATLAACAGATGCTGAGCAGACACAGACTTGGCTGCTGAGCTTCTGTTGGGAG 3420
OY 3421 ccagttgcccctcaattctcagaaggtatcgaattlaacgaatgcaactgtaacagtaaca 3480
DB 3421 CCAAGTTGCCCTCAATTTCTCAGAAGGTATCGATTTACGAATGCAACTGTGACAGTAACA 3480
OY 3481 aatattactgatatgaacacgttgaagttatctccaaagagaggttagagcagaacat 3540
DB 3481 AATATTACTGATGATGAATAACTGTTGAAGTTATTCTCAAAAGAGAGGTAGACGAGCAT 3540
OY 3541 gatgcaggtgctcactaagaagacatagtaattaaacaggttaactcctttagtactgat 3600
DB 3541 GATGCAAGGTGCTCACTAAGAAGACATTAAGTAATTAACACAGTAATCTTAGACTTGAT 3600
OY 3601 aacagcagacatataagattgttgaagttgaagttgaagatgagcagcaggtatgtagca 3660
DB 3601 AACAGCAGACATTAATTAAGATTGTTGTAAGTGAAGTTAAAGATGCAACAGTAATGTTGCA 3660
OY 3661 gatactattacattctataataagtaa 3687

```

```

DB 3661 GATACTATTACATTCATATTAAGTAA 3687
|||||
RESULT 3
AX033303
LOCUS AX033303 3687 bp DNA PAT 21-SEP-2000
DEFINITION Sequence 1 from Patent DE19903345.
ACCESSION AX033303
VERSION AX033303.1 GI:10280103
KEYWORDS
SOURCE Bacillus stearothermophilus.
ORGANISM Bacillus stearothermophilus
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 3687)
AUTHORS Lubitz W.
JOURNAL Patent: DE 19903345-A 1 03-AUG-2000;
LOBITZ WERNER (AT)
FEATURES
source location/Qualifiers
1..3687
/organism="Bacillus stearothermophilus"
/db_xref="taxon:1422".
sig_peptide
1..90
CDS 1..3687
1..3687
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/db_xref="GI:10280104".
/protein_id="CAC09982.1".
translation="MDRRKAVKLVATASA IAAAFYAAPNPNSAATDVATVYVQAKAO
FKKAVYVSHVYTERGEPNINDVYAEYNKAKRRYRDVAVLVNKGAKNDAYVADLO
KEYEYVFKARVKSQEARVAVYIDAVYAEKLEDRQLELAOVAQKDKEDQVYKI
PYEIKPTVYIDRYVGYKTTVDLRETFEFAKQLEDRQSLIYDLYVMAKREKQVAKG
NIDRAKAAVDQINQVLPKVDAKFKELTEVYKALDPAALATPVEVSVAINDONKA
VELTRVVPNGTILKQDLSAANEDVYVAVNTVYKIKYDQNIPLATRADVSLSDGRTIT
VDASTPEPNNTREYKRVKGIKDKNGKERKEAEFTKLNDVAVYVQFETNTNNSVN
LAAGTFEDDDLTVEVFKLAPLEYNSSNVTITDETGRKLPVLAISTGSTITTLIKE
ALVETGKQVKLAINNKTLLGYNAEVELEVPANSAAPVATAPPTLIGSTLSTLT
NMGKLAGVNEAGTVYVPGLOPTTTPARKRLDADNPEVYVERSGTVVASSELKYNA
DAKNTVLPKADIKENNTIYQIKIKKGLKSDGIEGTNYNKTPEPKTODITAPYISV
TSKNDAGIKATEQDEFTVRFSESNNTNATVSGSTITFGQVAVKAGANISATAS
DIIPASVEVAGQDSTYVYKVAANQLENDQKLVNKGATAPKADAMANTLATNY
IYTFPEQODVAPPTVYKFKGDSLKDAVATTLTVNDAGQKFTIOFSEBKTSGSL
VGKATVEKLTNNNGVADAGTGTVAATKTDANGVYAAVTLTGNDNDKAKRLV
VDSSTGDIADVAGNVIKEKQDILRYNSRHRVYASVYKAAADKQDQMSAARPTSAID
TKSLVLFNEDDLAEVREPNTVYVDAAGNVACTVYALDSTKPFVETPSDELKAGT
VYSITIDVNRDVGNTISKYITTSKTVSANPTLSISTADGVANDRSKTTITERSDS
VNPPTITLKKADGTSFTNTVTVANNENKTYKIFYHKVATLDEFTQYSLAVSKDFQTS
TDIDSKVTFIYGSVATDEVKPAALVGVSMGISTQDAAPATRLRVSADVFVAPALQF
SEGIDLTNAVTVYVITDDKTEVEVTSKESVDADHDACATKETIVINVTPLVDSKT
YKIVYSGVKAADGAVNDITTEYIK"
mat_peptide 91..3684
BASE COUNT 1316 a 660 c 715 g 996 t
ORIGIN
Query Match 100.0% Score 3687; DB 9; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 atcgaaagaaaagccttgaacatgacacagcagctgctatgacgaagtgcat 60
DB 1 ATCGAATAGAAAAGCCTTGAAACAGCAACAGCGCTATTCGACGAAGTCATTT 60
OY 61 gtcgcgcacaatccaaacgcttctgaagcgcgtacagatgtagcaacagtaagccaa 120
DB 61 GTCGCGCAATCCAAACGCTTCTGAAGCGGCTACAGATGTAGCAACAGTAGAACCCAA 120
OY 121 gcaaaagcagatcgaagaaacatactactacagcgaatacagtaacggaactggt 180
DB 121 GCAAAAAGCAGATTCAAAAGACATTAATTAATTCAGCCATACAGTAACGGAACCTGCT 180
OY 180
DB 180

```

QY	181	gaattcccaacatlacgcgtatcgtcgtacatacaaaagcgaaaaacgataccgt	240
DB	181	GAATTCCCAACATTATACGATGTATATGCGATATACAAAGCGAAAAACGATACCGT	240
QY	241	gctgcgntagcattagagaaataagcagcggcgagaaagacgccttactagtcgat	300
DB	241	GATGCGGTAAGATTAGTCAAATTAAGCAGCGCGCGCAAAAAGCGCTTACTACTAT	300
QY	301	ttacaaaaagataatgaacttaccgttttcaagcgaacccataactcggcgaagtcgt	360
DB	301	TTACAAAAAGATNTGAATACTTACGTTTTCAAGCAAAACCTAATNTGCGAACCTCGT	360
QY	361	gtcgcgaactacacgcgcgtctcaacactcgcgaacaaataagacgaatcgcgcgaag	420
DB	361	GTAGCAACTTCAATCGATCGATGCTTACACTAATGCAACAAAATTAAGCGAAATGGCGCAAG	420
QY	421	cttagcgctgctgtctcaagcaaaagttttagaagagcgaaacatactatccaaaatt	480
DB	421	CTTAGCGCTGCTGTCTCAAGCAAAAGATTTAGAAAAAGCAACATACTATCCAAATTT	480
QY	481	ccctatgaataataaacctgcacagtcattttagctcgcgtatctgylaaacacatcgt	540
DB	481	CCCTATGAATAATAAATCCGACAGTCATTTTGATCGCGTATNTGCTAAACACTCGT	540
QY	541	gatttactcgcctc-lacattlaaagcaaaagcaagaaacttcgcgcgaagcttaattat	600
DB	541	GATTTACTCGCTCTACATTATTAAGCAAAAGCACAGAACTTCGCGACGTTAATTTAT	600
QY	601	gattattccgctgtgaagcgcgcggagatccaaagacgcgtgtaagcggcgaattta	660
DB	601	GATATTACCGTGTGAAGCGCGCGAGATCCAAAGACGCGTGAAGCAGCGCAATTTA	660
QY	661	gacaaagctlaaagcttgcgtgttgcataacatcaactcacttccaaagtaacagatgt	720
DB	661	GACAAAGCTTAAAGCTGCGTGTGTGATCAAAATCAATACTTACCAAAAGTAAACATGCT	720
QY	721	ttcaaaactcgaactacagaagtcagcaaaaabgcatlga-lgcag-lgaagctgcgtc	780
DB	721	TTTCAAAACTGCACTTACAGAAAGTACGCAAAAAGCAATTAGATGCGAGATGAACTCCGCTT	780
QY	781	actccaagaagtgcgaag-lgtcaagtgcgattaaacacccaataaagcgttgtgaattaca	840
DB	781	ACTCCAAAGTGCAAAGTGTAACTGTAAGTCGATTAAACATCAAAACAAAGCTGTGAATTAACA	840
QY	841	gcagctaccagcggaaacgaaacataaattacaacttcacgttcgcgtgcgaatgaagataca	900
DB	841	GCACTACCGAGGAACGGAACACTTAAATTTACAACTTTCAGCTGCTGCACAAATGAATACA	900
QY	901	gtaaacgtaaatactcgtcagtcataclaaagtgcgcgtgaacatcccaattgccttaat	960
DB	901	GTAACGTAATACTGTAACGTATCTTAAAGTGAAGCGTAACATTCATTCGCCCTTAAT	960
QY	961	acgcgcagatgctttcttatctcacagcggaaaaactataactcgtgcgtctcaactcca	1020
DB	961	ACGCGCAAGTGTCTTCTTATCTTACAGACGCAAAAACATACGTGCACTCTTCAACTCCA	1020
QY	1021	ttcgaataataatcagcgatlaaagatagatagatgaagatgaatgaacaaaatgcgcga	1080
DB	1021	TTCCAAATTAATACGCGATTAAGATAGATGATTTAAAGGATTTAAAGACCAAAATGCGCAA	1080
QY	1081	gaattlaaagagaatgcattcactcttcgaagcttcgaaatgcagtcgttagtactcaagt	1140
DB	1081	GAATTAAAGAGATGCACTTCACCTTTCAGGCTTCGAAATGATGCTGATTACTCAAGTG	1140
QY	1141	tttggaaactaatgttaacaaacaaaccccttcgtcaactctgcagcagcgtlaacttgcac	1200
DB	1141	TTTGGAACTAATGTAAACAAACACACTTCTGTAAACTTAGTCAGCAGAGTACTTCCACACT	1200
QY	1201	gagcattccttaaaagatagatattgttgaatggtgtgcagcccgaaacgtaaaagcgcg	1260
DB	1201	GACATTAATTAACAGTAGATTTTGTGAATGATGTTTGAACCTGAACCTTAACAGCTCG	1260
QY	1261	aagcttccattacaagctgtgtgaactcgaaaacgcgattcccgtaattgcattcgtctct	1320

Db	1261	AACGTTACTATTACAGATGTTGAAACGCAAAACCATTTCCAGTAAATTCATCTACTCT	1
Qy	1321	ggtcttaacattactattagttaaanaagaggtttagtactcggfnaacaaataaaact	1
Db	1321	GGTTCTTACAAATTACTATTACGTTAAAGAAAGGCTTAGTAACTGGTAACAAATTAACCTT	1
Qy	1381	gctatcaataatgttaanaacattactcgtgttaacatgcagagcttaagagttagttc	1
Db	1381	GCTATCAATTAATGTTAAACCTTAACCGTTAACAAATGCAAGAAAGCTTACGAGTTAGTGTCC	1
Qy	1441	actgcaaaacgcaatgaagcaacaaactgtgctaccgctccbaactctttagtltgtaacat	1
Db	1441	ACTGCAAAACGATCAGACACCAACTGTGCTACCGCTCCATCTACTTTTAGTGCGTACAACT	1
Qy	1501	ttaattacvtgttctcttaacaaacaaatglttggfnaaatltgcgcggfngfngfnaagaa	1
Db	1501	TTATCTACTGGTCTCTTACAACAAAATGTTGGGTAATTTGGCTGGTGCTGAATGA	1
Qy	1561	gcttgaacttatactcgtgtcttcaatltcaacaaacaglttcgtactaaagtliagacgaa	1
Db	1561	GCTGGAACCTTTATTCGCGCTTCCTCAATTCACAAAGCTTTGGTCAATTAAGTTAGACGA	1
Qy	1621	tctacttagctgataaactltglatagltgnaaaagaactcgtgaacsgtltgttctct	1
Db	1621	TCTACTTAGTGGTGAATACTTGTGTTTAGTTCGAAGAAAGATCTGTAACAGTTGCTGTCT	1
Qy	1681	gaacataaataataatgcagacgcctcaaaaatgtaactttagtgcataaagcggaactaaa	1
Db	1681	GAACTTAAATTAATTAATGACAGACCTCAAAATGCTTAACCTTAGTCCCAAAAGCGACCTTAA	1
Qy	1741	gaataatacaatctatacaaalcaaaatlaaaaaagccttgaaagltccgaataaagtaagaa	1
Db	1741	GAAATTAACAACTTAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1
Qy	1801	ttaagcactgtttaaagcagaaaaacataatgattcaaaaactcaagacttaactgtctcaca	1
Db	1801	TTAGCCACTGTTTAACGAGAAACATATGATGTTCAAAATCAAGACTTAACGCTCCCTACA	1
Qy	1861	gttatctagctgataagcttaaaaatltgggcgcgtcgtgttlaaaagtacggaagccaagaa	1
Db	1861	GTTATTTACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTAACGTA	1
Qy	1921	tttactcgtgaagcttctccagagatlttaatacattlaatgcgtacacacagtttcggttagc	1
Db	1921	TTTTACTGGAAGTCTCAGAAATTTAAATACATTTAAATGCTACAAACGCTTCCGGTACC	1
Qy	1981	acaaacacaaatcaggttcaagtgtgctgtaglaaaacggtgtgcaacttactcgtccttaca	1
Db	1981	ACAAATCACATACGTCGAAGTGTCTGTAGTAAAGCGGTGCAAAATCTTAATCTGCTTACA	1
Qy	2041	gcaagttgacatcatctccagcttagtgttgaaagcggttactcgttcaagatgtgaacataaaa	1
Db	2041	GCAAGTTGACATCATTTCCAGCTAGCTGTTGAAGCGGTATACGTCAGAAATGCAACTACAAA	1
Qy	2101	gtgaaagttgtgtcttaacaaatlagaaacgtlaaaccaaggtatacaaatagtagttagtgcgt	1
Db	2101	GTTGAAAGTTGCTGCTAACCAATTAAGAACGTAAACCAAGGTTACAAATTAAGTAGTTCGCT	1
Qy	2161	aaaggtgcacacagcttccgtgtlaaagatgctgcgaatgtgcgaatlaactlttagcaactaat	1
Db	2161	AAAGGTGCACACGCTCGTTAAAGATGCTGCAAAATGCAAAATATCTTAGCACTAAGCTAT	1
Qy	2221	atctataatlaataaacttgaaggtlcaagacgttaacagcccaagcgttatacaaaagtatc	1
Db	2221	ATCTATATATTTAACAACTGAAGGCTCAGAGAGCTTAACAGCACCAACGCTTCAAAAGTTATTC	1
Qy	2281	aaagttgattctttaaanaagcgtctgtagcgttactacacacttagaagaagtttagtgcaggt	1
Db	2281	AAAGTGATTTCTTTAAAGAGCGCTGATGCACTTACTACACTTAACAAACTTGAATGCAAGGT	1
Qy	2341	caaaatcactatacaaatlttagcgaagaattaaaaactttagcgttctttagttaggt	1

[illegible][illegible]

1YTTTBCODVTAPEVTVKVFRODLSLKDAVATLTINWDAGKFTI0FSEELKTS6GL  
VGGKTVDEKLTNNNGVNDAGCTGTTSVAPRTDANGVTAAYVILITGDNDNDKAKLILV  
VKSSTIDGLADVANGVIEKEKDLIRYNSRHWVAAKAAADIGQADNAPPTSTALID  
TKKSLIVENEDLAIEVRENIVKDAANPLASTALDGTNKVFRPSLEKAGT  
VYSVTIDVROKDVNLTISKYISFKYVSAANPLASTALDGTNKVFRPSLEKAGT  
VYNTITLKKADGTSFTNTTLNINNNKNTYKIVHKGTTLDEFTOELAVSKRFDTG  
TDIDSKVFTIGSVATDEKPALEVGSGSNVSTYODAAATLRSVADVAAPVALCF  
SEGLDNTAIVTNTITDDKTVLEVISKESVDADHDAGAKETLINTVTPVLVDSKRT  
YKIYVSGVADAGNVADITTEYIK\*  
37.126  
sig\_peptide /gene="sbsa"  
mat\_peptide 127..3720  
/gene="sbsa"  
/product="surface layer protein"  
terminator 3767..3806  
BASE COUNT 1355 a 686 c 749 g 1042 t  
ORIGIN

Query Match 100.0%, Score 3687, DB 2; Length 3832;  
Best Local Similarity 100.0%, Pred. No. 0:  
Matches 3687: Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 1 atggataggaaaaaagctgtgnaactagcaacagcagtgctatlgcagcaagtgcatt 60  
DB 37 atggataggaaaaaagctgtgnaactagcaacagcagtgctatlgcagcaagtgcatt 96  
QY 61 gtgcgtgcaaatccaaagctcttgaaagcggtctacagatgttagcaacagtagtaagccaa 120  
DB 97 gtgcgtgcaaatccaaagctcttgaaagcggtctacagatgttagcaacagtagtaagccaa 156  
QY 121 gcaaaagcagcttcaaaaagcacaactactactacagccatacagttaacggaaactgct 180  
DB 157 gcaaaagcagcttcaaaaagcacaactactactacagccatacagttaacggaaactgct 216  
QY 181 gaattcccaacaatlaacgagtatagtctgaaatacacaagaacgaaaaaacsatccgt 240  
DB 217 gaattcccaacaatlaacgagtatagtctgaaatacacaagaacgaaaaaacsatccgt 276  
QY 241 ggtgcggtgagcattagtgaaagaagcaggtgagcggaagaagaagcgtactactagctgat 300  
DB 277 ggtgcggtgagcattagtgaaagaagcaggtgagcggaagaagaagcgtactactagctgat 336  
QY 301 ttacaaaaagaataatgaactacgctttcaagaacaaacccataatcctgcggaagctcgt 360  
DB 337 ttacaaaaagaataatgaactacgctttcaagaacaaacccataatcctgcggaagctcgt 366  
QY 361 gtacgaacttacaatcgatgtcttacaacta tgcacacaaaatagaagaaatgcgcaag 420  
DB 397 gtacgaacttacaatcgatgtcttacaacta tgcacacaaaatagaagaaatgcgcaag 456  
QY 421 ctgaagagctgcgtgltcaagcaaaaagatttagaaaagcagaacaatactatcacaatc 480  
DB 457 ctgaagagctgcgtgltcaagcaaaaagatttagaaaagcagaacaatactatcacaatc 516  
QY 481 ccttataaataaasctgcacagctcatttagatcgcgctataaggtaaaactcgt 540  
DB 517 ccttataaataaasctgcacagctcatttagatcgcgctataaggtaaaactcgt 576  
QY 541 gatttactcgtcttacaatttaangcaaaagcacaagaactcgcgcaacgcttaattat 600  
DB 577 gatttactcgtcttacaatttaangcaaaagcacaagaactcgcgcaacgcttaattat 636  
QY 601 gatttactcgtcttacaatttaangcaaaagcagaagctgcgtgaaagcagcaattta 660  
DB 637 gatttactcgtcttacaatttaangcaaaagcagaagctgcgtgaaagcagcaattta 696  
QY 661 gacaaagctaaagctgctctga tcaaatcaatcaatcttaccacaagaatgaagctgct 720  
DB 697 gacaaagctaaagctgctctga tcaaatcaatcaatcttaccacaagaagctgctgct 756  
QY 721 ttcaaaactgaactaagcagaagtaagcga aaaaagcattagatgcagatgaagctgct 780

DB 757 ttcaaaactgaactaagcagaagtaagcga aaaaagcatttagatgcagatgaagctgct 816  
QY 781 accccaaaagctgaagctgaaagctgcgtatcaactcaacaacaaaagctgttgaaattaca 840  
DB 817 actccaaaagctgaagctgaaagctgcgtatcaactcaacaacaaaagctgttgaaattaca 876  
QY 841 gcaatcagaatgaacggaacacaaaat tacaacttcagctgctgcaaatgaagatata 900  
DB 877 gcaatcagaatgaacggaacacaaaat tacaacttcagctgctgcaaatgaagatata 936  
QY 901 gtaacgtaaatactgttaactatcaataagctgagcaggtataacatcattgccttaac 960  
DB 937 gtaacgtaaatactgttaactatcaataagctgagcaggtataacatcattgccttaac 996  
QY 961 accgcaagctgttcttatactacagacggaacaaatacagctcgtggttcaactcca 1020  
DB 997 accgcaagctgttcttatactacagacggaacaaatacagctcgtggttcaactcca 1056  
QY 1021 ttcgaaataatacaggaataaagtagtagttaagtagttaaagataaaga caaaaatgcaaa 1080  
DB 1057 ttcgaaataatacaggaataaagtagtagttaagtagttaaagataaaga caaaaatgcaaa 1116  
QY 1081 gaattcaagaagaatgcaatcacttcaagcttcgaaatgaatgctgtagtlaactcaag 1140  
DB 1117 gaattcaagaagaatgcaatcacttcaagcttcgaaatgaatgctgtagtlaactcaag 1176  
QY 1141 ttgggaacttaattgaacaaacaaactctctgtaaaccttaacagcaggtacttccgaact 1200  
DB 1177 ttgggaacttaattgaacaaacaaactctctgtaaaccttaacagcaggtacttccgaact 1236  
QY 1201 gacgaatacttcaacagtagtaattgaaagctgttagcagctgaaactgtlaaacaagctg 1260  
DB 1237 gacgaatacttcaacagtagtaattgaaagctgttagcagctgaaactgtlaaacaagctg 1296  
QY 1261 aacgttacta tcaagaatgtgaacactggaacacgcatccagtaactgtactactctc 1320  
DB 1297 aacgttacta tcaagaatgtgaacactggaacacgcatccagtaactgtactactctc 1356  
QY 1321 ggttctacaattactatlaagttaaagaagaggttagtaactgttaacaaatataaact 1380  
DB 1357 ggttctacaattactatlaagttaaagaagaggttagtaactgttaacaaatataaact 1416  
QY 1381 gctacataatgttlaaaaacatlaactgtgttcaagaatgcaagaacttcaagcttaggtctc 1440  
DB 1417 gctacataatgttlaaaaacatlaactgtgttcaagaatgcaagaacttcaagcttaggtctc 1476  
QY 1441 actgcaaaagcaltcaagcaacaaactgtgtcactccactaactttagtgtgtacact 1500  
DB 1477 actgcaaaagcaltcaagcaacaaactgtgtcactccactaactttagtgtgtgtacact 1536  
QY 1501 ttactcaactggtctcttcaacaacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1560  
DB 1537 ttactcaactggtctcttcaacaacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1596  
QY 1561 gctggaacttactatcctgtgtcttcaacttcaacaacaaagcttgcataagtlagagcaa 1620  
DB 1597 gctggaacttactatcctgtgtcttcaacttcaacaacaaagcttgcataagtlagagcaa 1656  
QY 1621 tctacttagctgaataactttagttagtgaagaagaatcgtgtacaggtgtgtgtctc 1680  
DB 1657 tctacttagctgaataactttagttagtgaagaagaatcgtgtgtgtgtgtgtgtgtgtgt 1716  
QY 1681 gaactaaatataatgcaagcgttaaaatgtgttaactttagtgcacaaagcagcaattaaa 1740  
DB 1717 gaactaaatataatgcaagcgttaaaatgtgttaactttagtgcacaaagcagcaattaaa 1776  
QY 1741 gaaaatacaatcatcaaatcaaatcaaatcaaaaagcgtctgaagctcgaataagctatgaa 1800  
DB 1777 gaaaatacaatcatcaaatcaaatcaaatcaaaaagcgtctgaagctcgaataagctatgaa 1836  
QY 1801 ttaggcaactgttaaagagaacacata tgaattcaaaaactcaagacttaactgtcctaca 1860  
DB 1837 ttaggcaactgttaaagagaacacata tgaattcaaaaactcaagacttaactgtcctaca 1896

QY	1861	gtatattgagctaaagctttaaaatggcgcgcgtcgattaaagttacgtgaagcttaagaa	1970
Db	1897	gttt	1956
QY	1931	tttaccgtgaagttctccaagaatttaaatcaattaaagtcaacaacggtttcggtagc	1980
Db	1937	ttttactgtgaagttctccaagaatttttaatttaatttaatttaatttaatttaatttaattt	2016
QY	1961	acaatcacatacggctcaagttcgtctgattaaagcgggtgtcaactatctgctctaca	2040
Db	2017	acaattccacatacggctcaagttcgtctgattaaagcgggtgtcaactatctgctctaca	2076
QY	2041	gcaagtgaactcattccagctagctgtttgaagcgggttactggtcgaagatggaactataca	2100
Db	2077	gcaagtgaactcattccagctagctgtttgaagcgggttactggtcgaagatggaactataca	2136
QY	2101	gtgaagcttgctgtcaacccaattagaaacgtlaacccaagggtlaacaaatagtagttcgtgt	2160
Db	2137	gtgaagcttgctgtcaacccaattagaaacgtlaacccaagggtlaacaaatagtagttcgtgt	2196
QY	2161	aaagggtgaacaagctccgtttaagaatcgtctgaatatgaacatactttgacaactaatc	2220
Db	2197	aaagggtgaacaagctccgtttaagaatcgtctgaatatgaacatactttgacaactaatc	2256
QY	2221	atctatccatttaacattgaaggtccaagccgtaacagccaagccaagcggtttacaagatttc	2280
Db	2257	atctatccatttaacattgaaggtccaagccgtaacagccaagccaagcggtttacaagatttc	2316
QY	2281	aaagggtatctctttaaagaagcgtgtalgcagttacatacctacgaacgtttgacgt	2340
Db	2317	aaagggtatctctttaaagaagcgtgtalgcagttacatacctacgaacgtttgacgt	2376
QY	2341	caaaattccgactccaatttagcggaaattaaaaactcttagtgggtttctttagtggt	2400
Db	2377	caaaattccgactccaatttagcggaaattaaaaactcttagtgggtttctttagtggt	2436
QY	2401	ggccaagtaactgtgcgaagaaattlaacaacaacgaacgtagtggtagatgtcgtgatactga	2460
Db	2437	ggccaagtaactgtgcgaagaaattlaacaacaacgaacgtagtggtagatgtcgtgatactga	2496
QY	2461	actgtatcagttcgtcctctaaagaagaatgcgaatgtgtaagtaacacgtcgtgtgttaca	2520
Db	2497	actgtatcagttcgtcctctaaagaagaatgcgaatgtgtaagtaacacgtcgtgtgttaca	2556
QY	2521	ttaaatggtctctgcaatacgaacaagaatgcgaaatttgctgtcgttagtagataagct	2580
Db	2557	ttaaatggtctctgcaatacgaacaagaatgcgaaatttgctgtcgttagtagataagct	2616
QY	2581	tctatctatggaattgtctgtatgtagtgtgtaattgtaataaggaanaagataatttaatt	2640
Db	2617	tctatctatggaattgtctgtatgtagtgtgtaattgtaataaggaanaagataatttaatt	2676
QY	2641	cgttacaacagcctggagagacaacactgttagcttctgtgaaagctgcgtcgtgacaagaagtt	2700
Db	2677	cgttacaacagcctggagagacaacactgttagcttctgtgaaagctgcgtcgtgacaagaagtt	2736
QY	2701	caaaagcgttcctgtgatactccacaacaagaacgtgcaaatgtgatacaacaatgaagcttata	2760
Db	2737	caaaagcgttcctgtgatactccacaacaagaacgtgcaaatgtgatacaacaatgaagcttata	2796
QY	2761	gttgaattccaaltgaaactgtatttagcggaggttlaaacctgtgaacatcgttgttaagaat	2820
Db	2797	gttgaattccaaltgaaactgtatttagcggaggttlaaacctgtgaacatcgttgttaagaat	2856
QY	2821	gcagcaggttaatcgggttagctgtgatacgtcaacagcaattagacggtttctacaanaattt	2880
Db	2857	gcagcaggttaatcgggttagctgtgatacgtcaacagcaattagacggtttctacaanaattt	2916
QY	2881	gtattcaactcaattccaagaattaaagcgttgtaagttactctgtgtacacatggaaggt	2940
Db	2917	gtattcaactcaattccaagaattaaagcgttgtaagttactctgtgtacacatggaaggt	2976

QY	2941	gttgagatataaagtagglaacacaalcctcaaaiaaattacttggttaagaacgttactc	3000		
Db	2977	gtgagagatgtaaaagtaggtaaacaaactctcaaaatgacattgcttccaaagcctgactc	3036		
QY	3001	gcgaatccaaagtgatctctccalcacgaatgctgcgaagctgcgaagcttaacglttaacgcttct	3060		
Db	3037	gcgaatccaaagtgatctctcaatcagatgctgcgaagctgcgaagcttaacglttaacgcttct	3096		
QY	3061	aaacaattacaaattgaattcagcagatcagttccaaaaccaacaatcactcttaagaag	3120		
Db	3097	aaaacattgtaaaatgaaatgaaacagatgctgcgaagctgcgaagcttaacglttaacgcttct	3156		
QY	3121	gctgaagcaacttactatcttaatacttaacttgatgaatgtaataataaataaataaaca	3180		
Db	3157	gctgaagcaacttactatcttaatacttaacttgatgaatgtaataataaataaataaaca	3216		
QY	3181	tacaagaattgatacccaagaagtgacaacactgcaagcttactccaalatagatgaaca	3240		
Db	3217	tacaaatgtaattgattcccaaaagctgtaacactgcaagcttactccaalatagatgaaca	3276		
QY	3241	gttccaaaagaatttccaaactggttaactgatatgtagaagaaagttaacttaacagat	3300		
Db	3277	gtttcaaaaagatttccaaactggttaactgatatgtagaagaaagttaacttaacagat	3336		
QY	3301	tcgtgtcctcctgcgaaggaataaacctgctctatgttggttggttccatgaaatgaaaca	3360		
Db	3337	tcgtgtcctcctgcgaaggaataaacctgctctatgttggttggttccatgaaatgaaaca	3396		
QY	3361	agctatactcagaatgtctgcgaagaacaacgacttcgaactgtaagctgaactcctcttcgag	3420		
Db	3397	agctatactcagaatgtctgcgaagaacaacgacttcgaactgtaagctgaactcctcttcgag	3456		
QY	3421	ccgaattgcctctcaattctcagaaggtatcgaatttaagaatgcaactgctgcgaactaaca	3480		
Db	3457	ccgaattgcctctcaattctcagaaggtatcgaatttaagaatgcaactgctgcgaactaaca	3516		
QY	3481	aacttactggttgataaactggttggaagttaatttcaaaagaaaggtgtaagcaagacat	3540		
Db	3517	aacttactggttgataaactggttggaagttaatttcaaaagaaaggtgtaagcaagacat	3576		
QY	3541	gaatgcaggtgctataagaagacaatgaataataacacagttactcctcttaagtaattgat	3600		
Db	3577	gatgcaggtgctataagaagacaatgaataataacacagttactcctcttaagtaattgat	3636		
QY	3601	aacagcaagacttaaaagattgtgtgaagtggaagtaaagaatgcagcagtaagtgtga	3660		
Db	3637	aacagcaagacttaaaagattgtgtgaagtggaagtaaagaatgcagcagcagtaagtgtga	3696		
QY	3661	gatcattatcattctatataatga	3687		
Db	3697	gatcattatcattctatataatga	3723		
RESULT 5					
LOCUS	AF055578	3680 bp	DNA		
DEFINITION	Bacillus stearothermophilus surface layer protein precursor (sbsc)				
ACCESSION	AF055578	gene, complete cds.			
VERSION	AF055578.2	GI:4581955			
KEYWORDS					
ORGANISM	Bacillus stearothermophilus.				
SOURCE	Bacillus stearothermophilus				
REFERENCE	1 (bases 1 to 3680)				
AUTHORS	Jarosch,M., Egelsei,E.M., Matdanovich,D., Sleytr,U.B. and Sara,M.				
TITLE	S-layer gene sbsc of Bacillus stearothermophilus ATCC 12980:				
	molecular characterization and heterologous expression in				
	Bacillus coli				
JOURNAL	Microbiology 146 (Pt 2), 273-281 (2000)				
MEDLINE	20170659				
PIRBASE	10708365				



```

Db      1344  AAGCAGTACAGCTTCATTACATTCAACCTTCATTACAGTACCTGACATGATGAGCTT 1403
QY      1141  ttggaactaa 1151
                |
                |
                |
Db      1404  GATTTGACTAA 1414

RESULT   6
LOCUS    AF228338              3090 bp      DNA      BCT      24-JUL-2000
DEFINITION Bacillus stearothermophilus Sbsd precursor (sbsd) gene, complete cds.
ACCESSION AF228338
VERSION   AF228338.2
KEYWORDS  GI:9186638
SOURCE    ORGANISM
          Bacillus stearothermophilus.
          Bacillus stearothermophilus
          Bacteria; Firmicutes; Bacillus/Clostridium group;
          Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 3090)
AUTHORS  Danhorn,T., Egelseer,E.M., Idtiss,R., Sleytr,U.B. and Sara,M.
TITLES   Molecular characterization of the first bacterial S-layer
          glycoprotein Sbsd from a temperature derived mutant of Bacillus
          stearothermophilus ATCC 12980
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 3090)
AUTHORS  Danhorn,T., Egelseer,E.M., Idtiss,R., Sleytr,U.B. and Sara,M.
TITLES   Direct Submission
          Submitted (24-JAN-2000) Zentrum fuer Ultrastrukturforschung,
          Universitaet fuer Bodenkultur, Gregor Mendel-Strasse, Vienna
          A-1190, Austria
          3 (bases 1 to 3090)
          Danhorn,T., Egelseer,E.M., Idtiss,R., Sleytr,U.B. and Sara,M.
          Direct Submission
          Submitted (14-JUL-2000) Zentrum fuer Ultrastrukturforschung,
          Universitaet fuer Bodenkultur, Gregor Mendel-Strasse, Vienna
          A-1190, Austria
REMARK    Sequence update by submitter
COMMENT   On Jul 14, 2000 this sequence version replaced gi:6984125.
FEATURES
          location/Qualifiers
          1..3090
          /organism="Bacillus stearothermophilus"
          /db_xref="taxon:1422"
          255..261
          /gene="sbsd"
          273..2984
          /gene="sbsd"
          273..2984
          /gene="sbsd"
          /note="surface layer glycoprotein"
          /codon_start=1
          /transl_table=1
          /product="Sbsd precursor"
          /protein_id="AA04763.2"
          /db_xref="GI:9186638"
          /translation="MDKKKKKATATAGAAASAVVAANPHASQAATDVATVYVSQAKAK
          MKKAYTISHRVETGKLPNSDYIAATNNAKQAYAAVAAVNAGAKKAYIADILQ
          AYETVYFKNPKSGEARVATYIDAVYATKIDANRELKAAVEAKDKKAEELYHKIK
          SYEKTRVILIDRYVCGSTREILRSTFKADQALRDSITIDITVAMKAREQDAKAGAK
          NDKRAALDOVNOYVSGVNDPAKAEELQKAQDANNAVEALPKVESVTVNAKTTLE
          LKPKKADYAVNVDNKGTSDDSDVVKAKATILTRIDGGCGSVYKASLSDDKTKL
          VADDSOFTRKYVVDIKNKVTLDSKDVPSITTTIDTDSVBPVSLSSYADNGILTVV
          KESPLASVGVIKLDGTTELSSVSPKFTAGDETTINLASSVYVNDLTLEFGADV
          YNGVAVNPINPAELVTKETVDITPKPVQNIETAVNTKVKVFSERKLKGGQGT
          ASVSDTGLVYATLANALSEGVYAVEVSDYKDLQAGAGDITGVQNLKDTAPKRE
          VSGVYVLDIEVEHVLPRDEVTGSMITVQASDKVDEYNNYKISVGTLLTTSDNEF
          KLYLPFGKSKSAVALNISLSPKGYTYTLPDRLEVSLDAGNPKYERKQITTPKGSDEL
          TRKALDSSYDNGVGAADNNELVFAEFOQNDASALNSTRINGLAVTKAVFDGDR
          IRTVLGAGANTYGTGHTTISNKNITGLVMDTVTSEFKENINLAPTETALISADYI
          RVFSEPVANATISSPLANNFTVKVCGNVATVENVYEDNNATQVGGSGYKTVYLLK
          LOSGVSLSKRPITLSAGIYDVDTGALTNNNLNGNNVSNAAVYAK"
          273..362
          /gene="sbsd"
          3010..3050
          terminator

```

BASE COUNT	941 a	716 c	746 g	687 t
ORIGIN				
Query Match	13.2% Score 487.6; DB 2; Length 3090;			
Best Local Similarity	73.9%; Pred. No. 1.8e-74;			
Matches 619; Conservative	0; Mismatches 219; Indels 0; Gaps 0			
QY	1	atggaatgaaaaaagctgttgaaactagcaacagcaagctgtctatltgcagcaagtgcattt	60	
Db	273	ATGGACAAGAAAGAAAGCAAGTCAAACTCCGCCACAGCCAGCGTGTACCAAGCAAGCGCTTTC	332	
QY	61	gtcgctgtcaaatccaacagcttctgaagcgtgtacagatgtagcaacagtagaaaccaa	120	
Db	333	GTCCGCGCCCAACCCGACGCTTCCCAAGAGCAACCGAGGTGGCGACGCTGTCTCACGCCAA	392	
QY	121	gcaaaagacacagcttcanaaaagcatatactactactaacgcatacagataaagcaagacttgt	180	
Db	393	GGGAAGACGCCCAATTAAGAAAGCAACATTAATCTATACGTACAGCCACAGCGTAAACGGAAAGCGGC	452	
QY	181	gaattccaaacattaaacagatgtatattgtctgtaatacaacaagaacgaaaaaacgatacagt	240	
Db	453	AAATCCCGAATCATGATGACGTATATACGCGCGCTACAAACAGCGCAAAACAAAGCGTATGCG	512	
QY	241	gattcggtgacatttagtaataaagcaggtgacgcgcaaaaaaagacgcttaacttaagctgat	300	
Db	513	AACGGCGTCGCGTGTGTGTGAACAAACCCGCGCGCCGCAAAAAGACCGGATTTGGCGTGAT	572	
QY	301	ttacaaaagaatattgaacattacgcttcttcaagaacaaaccttaactctgacagctcgt	360	
Db	573	TTGCAAGCTACTATACGAAACGTAATGATTATTCAAAGCCAAACCCGAAATTCGGCGAAGCGCGC	632	
QY	361	gtagcaacttlaactgcattgtcttacaactatgcaacaaatattgaacgaaatgtgcgcaagag	420	
Db	633	GTCCGCAAGCTACATTGAGCGCTTACAACTATGCGACAAATTTAGACCGCATTCGCTGAGGAA	692	
QY	421	ctagagctgtctgtctcaagaacaaagatttagaaaaagcaacaaatactatacaaaact	480	
Db	693	CTCAAAAGCAAGCGGTGCAAGCGCAAGATCTGAAAGAAAGAGAGCTCTACCAAAATTT	752	
QY	481	ncctttagaattaaacacgcacagcucatttttagatcgggtatattgttaaacacacttgt	540	
Db	753	TGCTATGAGCTTAANAACGGCACGGGTCATTTTCGACCGCGCTCTACGCCAATGTGACGCCG	812	
QY	541	gatttacttcgctctcaatttaaagaacaaagcaacgaacttcgcgcacagcttaattat	600	
Db	813	GAACTCCCTCCGTTGCAAGCGTCAAAAGCGATGCTCAAGCATTTGCGGACAGCTGTGATCTAC	872	
QY	601	gatataccgctgtcaacatgaagcgcgcggaagtaacaagacgcgtgtgaagcaggaattta	660	
Db	873	GACACACACTGTTGCCATTAAGAACCGGTGAAGCGCAAGAGCGCGCCAAACCCGCCAACAACCTG	932	
QY	661	gacaaagctaaagctgcgcgtgtgataaatcaatcaatctacttccaaagaataagatgt	720	
Db	933	GACAAAGGAAAGCGCGCTTGACCAACGTCAACCAATTAATGTAATCGAAATTAACCGAATGG	992	
QY	721	ttcaaaactgaactacacaggaatgcgaaaaaagcatltagatgcgaattaaagctgcctt	780	
Db	993	TTCAAAAGCGCAACTCCAAAAAGCAACGCGCAAGACCCCAACCGCGCGGTATGTAAGCGCGGTTG	1052	
QY	781	actccaaagatgtgaagtgtaagtgcatltaacactcaactaaacaaagctgtttgaattaa	838	
Db	1053	CTTCCAAAGATGTGAAGTGTATACACCGCGTTAACGCCAATAAGCTGTGAATTCATAATTC	1110	
RESULT	7			
ID	E10125	standard; RNA; UNC; 2069 BP.		
XX	E10125;			
XX	E10125.1			
SV				





Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://greml1.zool.iastate.edu/cgi-bin/sp.cgi>). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is T8N9 and the 3' clone is T7B9.

FEATURES  
source

1..111554

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db\_xref="taxon:3702"

/chromosome="3"

/clone="F1D9"

/clone\_lib="IGF BAC"

/complement(join(4747..5195,5297..5754,5827..6284,6457..6803,6864..7079,7216..7286,7326..7810,7914..8253,8311..8737,9083..9347))

/note="gene\_id:F1D9.1"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/product="helicase-like protein"

/join(11589..11736,11836..12019,12305..12497,12774..12896,13021..13080)

/note="gene\_id:F1D9.2"

/codon\_start=1

/evidence=not\_experimental

/product="ATP-dependent CLPB protein-like"

/protein\_id="BA97091.1"

/db\_xref="GI:8777573"

/translation="MEEDLHERVTAQDLDESVAQDAICSKAGISDPNPIASENEMGPFSVQAVGASVGVYEGDKSTEVRRPVSVQDEIEKPPNPFSLQLDDEPLTDSHSLILETIRNEDIEAVYEMAKQCVELARTEPKFENRIDEIVSGPLNSVSKIVELOVVOGRKKRLEQPNNTLEITLVNLRKDPFENDTILIDADQPEKLVTKKNNNAVNK"

/complement(join(19018..19216,19317..19390,19600..19746))

/note="contains similarity to unknown protein"

/gene\_id:F1D9.3"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="BA97092.1"

/db\_xref="GI:8777574"

/translation="NAKKERHRVGVREEDVVRKMLINPYLINESLHOLRALFSGYPSTTLITVYMKLEEMKMTKQEFIASFRSSGFTGAYTPMCAKINNDGKGLAVLPQIRITLETLYSKSKLGIKKDLITIVIVYKLVMP"

/join(21817..22004,22075..22143)

/note="contains similarity to ATP-dependent CLPB protein"

/gene\_id:F1D9.4"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/complement(22489..23812)

/note="contains similarity to retroelement pol

polyprotein

/gene\_id:F1D9.5"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/join(24404..24543,24659..24879)

/note="contains similarity to ATP-dependent CLPB protein"

/gene\_id:F1D9.6"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/complement(25073..29379)

/note="gene\_id:F1D9.7"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/product="retroelement pol polyprotein-like"

/join(30097..30190,30265..30551,30614..31510,31758..32036)

/note="gene\_id:F1D9.8"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/product="ATP-dependent CLPB protein-like"

/join(40639..40982,41779..43326,42993..42794,42903..42938,43020..43142,43226..43353,43425..43481,43588..43650,43733..43973,44026..44107,44503..44574,44813..44940,45043..45198)

/note="gp|AAD25852.1"

/gene\_id:F1D9.9

/strong similarity to unknown protein"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/join(46889..47780,47984..48332,48695..48783)

/note="gene\_id:F1D9.10

/pir||T01475

similar to unknown protein"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/join(5111..51234,51470..52398,52474..52672,52762..53360,53433..53558,53717..53809,53908..54010,54110..54210,54472..54531,54627..54728,54795..55047,55142..55226,55334..55637,57320..57389,57601..58187)

/note="gene\_id:F1D9.11

/pir||T02599

strong similarity to unknown protein"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/complement(join(58782..58995,59066..59150,59231..59498,60276..61922,61970..62259,62305..62482))

/note="gene\_id:F1D9.12"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/product="maize transposon MuDR protein-like"

/complement(join(63866..64008,64160..64226,64340..64414,64504..64620,64743..64790,64956..65053,65183..65264,65486..65566,65677..65718,65834..65952,66106..66144))

/note="contains similarity to human diaphanous 1"

/gene\_id:F1D9.13"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="BA97094.1"

/db\_xref="GI:9294133"

/translation="MDSKVIPLPDMSVTLALDESVDVDVNDLIKPCPTKEAEILKGFIGNKETTAKCEQFLELTKPVEVEKLVESKIQFHSQVRSSTLTKRIMQITLSIGNALNHCITARGSAIGPRIDSLIKTIDRSNSKSTLHNYLCVLAKEKLGILNFPKMDIOLKTYLAEMQATSKGLERVOGFTASTDQCSKTFHNLKEFLSVAGEVYSTASLGVGGSADALALYFGEDEPARVDFQVSTLQNFVAIFRSHENCKQVEFEKRAKFAENKLRKGVYNN"

/complement(66524..66974)

/note="gp|AAD23008.1"

/gene\_id:F1D9.14

similar to unknown protein"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/complement(join(68701..69561,69697..70591,70640..70667,71526..72709,72898..73300,73370..73533,73687..74594))

/note="gene\_id:F1D9.15"

/codon\_start=1

/pseudo

/evidence=not\_experimental

/product="retroelement pol polyprotein-like"

/complement(75178..76058)

/note="gp|AAD23008.1"

/gene\_id:F1D9.16

```

similar to unknown protein"
/codon_start-1
/pseudo
/evidence=not_experimental
76515..80547
/note="emb|CA875469.1
gene_id:RID9.17"
/codon_start-1
/pseudo
/evidence=not_experimental
/product="copia-type reverse transcriptase-like protein"
complement(join(81294..81521,81620..82285))
/note="gb|AAP23008.1
gene_id:RID9.18
strong similarity to unknown protein"
/codon_start-1
/pseudo
/evidence=not_experimental
complement(join(87685..88025,88101..88185,88231..88862,
88953..89075,89158..89715))
/note="gene_id:RID9.19"

```

```

similar to unknown protein"
/codon_start-1
/pseudo
/evidence-not_experimental
76515..80547
/ncbi="emb|CA875469.1
gene_id:FLI09.17"
/codon_start-1
/pseudo
/evidence-not_experimental
/product="copa-type reverse transcriptase-like protein"
complement(join(11294..81521,81620..8285))
/ncbi="gb|AA053008.1
gene_id:FLI09.18
strong similarity to unknown protein"
/codon_start-1
/pseudo
/evidence-not_experimental
complement(join(87665..88025,88101..88185,88231..88862
88953..89075,89158..89775))
/ncbi="gene_id:FLI09.19"

```

Query Match	2.1%	Score 77	DR 13	Length 111554
Best Local Similarity	44.3%	Pred. NO. 0.00051		
Best Match 314, Conservative	0	Mismatches 395	Indels 0	Gaps 0

[illegible]

RESULT 10  
 AF056936 LOCUS 5420 bp DNA INV 13-APR-1998  
 DEFINITION Plasmodium falciparum mature parasite-infected erythrocyte surface antigen gene, complete cds.  
 ACCESSION AF056936  
 VERSION AF056936  
 KEYWORDS GI:3044184  
 SOURCE  
 ORGANISM  
 'malaria parasite P. falciparum.  
 Plasmodium falciparum  
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 REFERENCE 1 (bases 1 to 5420)  
 Coppel, R.L.  
 Repeat structures in a Plasmodium falciparum protein (MESA) that binds human erythrocyte protein 4.1  
 Mol. Biochem. Parasitol. 50 (2), 335-347 (1992)  
 JOURNAL  
 MEDLINE 92158014  
 PUBMED 1741020  
 2 (bases 1 to 5420)  
 Kun, J.F.J., Waller, K.L. and Coppel, R.L.  
 Plasmodium falciparum: structural and functional domains of the mature-parasite-infected erythrocyte surface antigen (MESA/PfEMP2)  
 Unpublished  
 3 (bases 1 to 5420)  
 Kun, J.F.J. and Coppel, R.L.  
 Direct Submission  
 Submitted (01-APR-1998) Humanparasitology, Institute for Tropical Medicine, Wilhelmstrasse 27, Tübingen 72074, Germany  
 FEATURES  
 source location/Qualifiers  
 1..5420

[illegible]

CD5

mRNA

```

/organism="Plasmodium falciparum"
/strain="Palo Alto"
/db_xref="taxon:5833"
join(<1..150,281..>:5116)
/product="mature parasite-infected erythrocyte surface antigen"
join(1..150,281..5116)
/note="MESA, PfEMP2"
/codon_start=1
/product="mature parasite-infected erythrocyte surface antigen"
/protein_id="PAC13303.1"
/db_xref="GI:3044185"
/translation="MEVIGRNLGYDRKNNMWDGNGNVKVVNNSSLRKYKFCQCTI
ICVGLDITNGCSPTYSYSIMNNNDYRIILSTPMPSSRLTEIMTPDIEDLSIR
NYITCLRNAPYIDDPIMGSYVYTDKRNNDQYIKLIMDQESRRKQKQBNADIEI
RKKEKELAKELMDSDVDKARFIQFIKLDRTPRNVAVHVMGSGESKEDISKEI
ESKTKGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKET
TGSEKTKGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETVY
KITHETPSTENVKLTTESKDREGNVKAGPENSNSVTSSETPKTLAKBENEGK
LGVNVTGASGNSNDPDKLTIDQENKNGKVSPPSPEDDKPENNKKKSKKKKKS
PFGQLGCPFLCNRKITETDDEBETLVYADKAKKHKPLRANTERKNDNEKKDLIGSD
KEDYKNEEDQKDVLDGSDQEDKVKKNDQKVLGSDQEDVKENDKQKQVIGS
EKQDTEKREKVRKVRKRRKRVKRGKJENDTBGNDKVPGLIIEBVKELIKRGSD
GJENDTGNDKYGPGLIIEBVKELIIRKNSVETGNDKYKGPGLIIEBVKELI
EIKHVEBEGIKENDTESKDVIGOEILTEBVBEGIKENDENDKYGPGLITEBVK
EIBQETKGRKNLEIKDVLIGQVLIIEBVKVIRKKVYKJGKJKNHTESKDVIGOE
LIEBVAEELIKQVEBEGIKENDTESKDVIGOEYKJGVABEGPNDKDTQKQKVAE
VKEBVAEELIKRVKRRKNNKERNQDNYGKELIMEDVNERKDTANKDEITQEKREBVK
EKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREB
IEQKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKRE
EINNMERKDTANKDVIIEQKREBVKREBVKREBVKREBVKREBVKREBVKREB
KDEITQEKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKRE
VIGVIMNEDVNEKDTANKDVIIEQKREBVKREBVKREBVKREBVKREBVKREBVK
DTEIDQEKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKREBVKRE
VIGVIMNEDVNEKDTANKDVIIEQKREBVKREBVKREBVKREBVKREBVKREBVK
VKEBVAEELIKRVKRRKNNKERNQDNYGKELIMEDVNERKDTANKDVIIEBVKELI
KRVKVRKRRKNNKERNQDNYGKELIMEDVNEKDTANKDVIIEBVKELIIEBVAE
VYKVRKVRKRRKNNKERNQDNYGKELIMEDVNEKDTANKDVIIEBVKELIIEBVAE
QEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBEBE
ESBEBKVTSGVVAHRMCKRKKRKNNGLEIYVAKQNEBVFDEBYIIGQVIMTLPF
PANAIDDEPRFPHNLIIEHDNDMLKLIYNIIEYVPHANNKQKELDKPESFM

```





OY	2859	agacgctgtacaataaaatttgcttcttcaccatcccaagaagataaaagcgggnaagt	2918
DB	175973	TGATTTCCTATTTCTTGCGCTAGTATGTGCCTGGTGGCTTAAATTGGAAGGATGT	176032
OY	2919	ttaactcgttaacnattgcacggtgtgaagatagaagtaagtaaacacaaatccatacat	2978
DB	176033	TGATTTCCTGTTGCTAATGGTGGCTGACATATTTGCTAAATTTGGTGGCTGTGTGCAATTTGT	176092
OY	2979	tacttcglttaagaactgtatctcgcgaatccaacgcttatcttcaalcaagcattgcgcgaag	3038
DB	176093	TGGTTTGTGTGTTACTGGTTTTCGGTTTGGCTTTTGTAATGTTGGTGGCTGGCTGGTGGCTGGTGC	176152
OY	3039	tgcagttaacgttcacgcgcttccaaaacatacaatt	3075
DB	176153	TGCTGCTGCTGATGCTGCTGTTACTGTTACTATTTCTGATT	176189
RESULT	12		
LOCUS	AC087877/c		
DEFINITION	AC087877	mus musculus clone RP23-9L21, ** SEQUENCING IN PROGRESS **	32
ACCESSION	AC087877		
VERSION	AC087877.2	GI:13122756	
KEYWORDS	HTG, HTGS, PHASEI.		
SOURCE	house mouse.		
ORGANISM	Mus musculus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1 (bases 1 to 208006)		
TITLE	McCombie,W.R., Baker,J.P., Bahuet,A., Bal,H., Bailja,V.,		
REFERENCE	Dedhia,N.N., de la Bastide,M., Huang,E.N., King,L., Kirchoff,K.A.,		
AUTHORS	Miller,B., Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R.,		
JOURNAL	Rodriguez,S., Santos,L., Shah,R.S., Spiegel,L.A., Toth,K., Vil,M.D.		
TITLE	and Zintaverni,T.		
REFERENCE	Mouse Genomic Sequence		
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 208006)		
TITLE	McCombie,W.R.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (03-FEB-2001) Lita Annenberg Hazen Genome Sequencing		
JOURNAL	Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring		
TITLE	Harbor, NY 11724, USA		
REFERENCE	On Feb 24, 2001 this sequence version replaced gi:12658030.		
COMMENT	----- Genome Center		
	Center: Lita Annenberg Hazen genome Center, Cold Spring Harbor		
	Laboratory		
	Center code: CSHL		
	Web site: http://www.cshl.org/genseq		
	Contact: mccombie@cshl.org		
	----- Project Information		
	Center project name: RP23-9L21		
	Center clone name: RP23-9L21		
*	* NOTE: This is a 'working draft' sequence. It currently		
*	* consists of 32 contigs. The true order of the pieces		
*	* is not known and their order in this sequence record is		
*	* arbitrary. Gaps between the contigs are represented as		
*	* runs of N, but the exact sizes of the gaps are unknown.		
*	* This record will be updated with the finished sequence,		
*	* as soon as it is available and the accession number will		
*	* be preserved.		
*	I	23097:	contig of 23097 bp in length
*	23098	23329:	gap of unknown length
*	23330	38329:	contig of 16000 bp in length
*	39330	39561:	gap of unknown length
*	39562	53028:	contig of 13467 bp in length
*	53029	53260:	gap of unknown length
*	53261	64737:	contig of 11477 bp in length
*	64738	64969:	gap of unknown length
*	64970	74807:	contig of 9838 bp in length
*	74808	75039:	gap of unknown length

	*	75040	84783:	contig of 9744 bp in length
	*	84784	85015:	gap of unknown length
	*	85016	94437:	contig of 9442 bp in length
	*	94438	94669:	gap of unknown length
	*	94670	103283:	contig of 8614 bp in length
	*	103284	103515:	gap of unknown length
	*	103516	110591:	contig of 7576 bp in length
	*	111092	111323:	gap of unknown length
	*	111324	118475:	contig of 7152 bp in length
	*	118476	124866:	contig of 6160 bp in length
	*	118707	125097:	gap of unknown length
	*	124867	131107:	contig of 6010 bp in length
	*	131108	131338:	gap of unknown length
	*	131339	137094:	contig of 5756 bp in length
	*	137095	137325:	gap of unknown length
	*	137326	142821:	contig of 5486 bp in length
	*	142822	143052:	gap of unknown length
	*	143053	148485:	contig of 5433 bp in length
	*	148486	148716:	gap of unknown length
	*	148717	154023:	contig of 5307 bp in length
	*	154024	154254:	gap of unknown length
	*	154255	159032:	contig of 4778 bp in length
	*	159033	159263:	gap of unknown length
	*	159264	164021:	contig of 4758 bp in length
	*	164022	164252:	gap of unknown length
	*	164253	168884:	contig of 4632 bp in length
	*	168885	169115:	gap of unknown length
	*	169116	173323:	contig of 4208 bp in length
	*	173324	173554:	gap of unknown length
	*	173555	177675:	contig of 4121 bp in length
	*	177676	177906:	gap of unknown length
	*	177907	181802:	contig of 3896 bp in length
	*	181803	182033:	gap of unknown length
	*	182034	185114:	contig of 3081 bp in length
	*	185115	185345:	gap of unknown length
	*	185346	188336:	contig of 2991 bp in length
	*	188337	188567:	gap of unknown length
	*	188568	191342:	contig of 2775 bp in length
	*	191343	191573:	gap of unknown length
	*	191574	194263:	contig of 2630 bp in length
	*	194264	194494:	gap of unknown length
	*	194495	197136:	contig of 2642 bp in length
	*	197137	197367:	gap of unknown length
	*	197368	199977:	contig of 2610 bp in length
	*	199978	200308:	gap of unknown length
	*	200209	202750:	contig of 2542 bp in length
	*	202751	202981:	gap of unknown length
	*	202982	205344:	contig of 2353 bp in length
	*	205345	205575:	gap of unknown length
	*	205576	207749:	contig of 2174 bp in length
	*	207750	207980:	gap of unknown length
	*	207981	208006:	contig of 26 bp in length.

FEATURES  
source Location/Qualifiers

```

1..208006
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone="RP23-8i21"

```

BASE COUNT 59324 a 40597 c 40553 g 60335 t 7197 others

ORIGIN

Query Match 1.8% Score 66; DA 77; Length 208006;  
Best Local Similarity 46.4%; Pred. No. 0.037; Indels 0; Gaps 0;  
Matches 216; Conservative 0; Mismatches 250;

DY 11 aaaagctgtgaactatgcaacagaagtgcatttcgacgaagtcatttgcgtcaca 70  
||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 31578 AAGAGGAAGGAGGAGAGAGACGAGAGAAAGCAGGAGAGACAGACAGAGAAGAGCG 31519  
||| | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 71 atccaaagccttcgaagcgagctcacagatgtagaacagtagtaagccaagaagaagcac 130  
||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 31518 AAGCGAAGAGAGCACAG 31459  
||| | | | | | | | | | | | | | | | | | | | | | | | | | |









**THIS PAGE BLANK (USPTO)**